

RAW SEQUENCE LISTING  
PATENT APPLICATION US/09/492,697DATE: 02/16/2000  
TIME: 02:14:57

INPUT SET: S34756.raw

This Raw Listing contains the General  
Information Section and up to the first 5 pages.

MAY 01 2000  
TECH CENTER 1600/2900

## SEQUENCE LISTING

ENTERED

## (1) General Information:

(i) APPLICANT: Choulika, Andre  
Perrin, Arnaud  
Dujon, Bernard  
Nicolas, Jean-Francois

(ii) TITLE OF INVENTION: Nucleotide Sequence Encoding the Enzyme  
I-SCEI and the Uses Thereof

(iii) NUMBER OF SEQUENCES: 52

## (iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &  
Dunner  
(B) STREET: 1300 I Street, N.W.  
(C) CITY: Washington  
(D) STATE: D.C.  
(E) COUNTRY: USA  
(F) ZIP: 20005-3315

## (v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Floppy disk  
(B) COMPUTER: IBM PC compatible  
(C) OPERATING SYSTEM: PC-DOS/MS-DOS  
(D) SOFTWARE: PatentIn Release #1.0, Version #1.25

## (vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER: Unassigned  
(B) FILING DATE: 25-JAN-2000  
(C) CLASSIFICATION:

## (vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: US 09/244,130  
(B) FILING DATE: 04-FEB-1999

## (vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: US 09/119,024  
(B) FILING DATE: 20-JUL-1998

## (vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: US 08/336,241  
(B) FILING DATE: 07-NOV-1994

RAW SEQUENCE LISTING  
PATENT APPLICATION US/09/492,697MAY 01 2000  
TECH CENTER 1600/2900  
DATE: 02/16/2000  
TIME: 02:14:57

INPUT SET: S34756.raw

47 (vii) PRIOR APPLICATION DATA:  
48 (A) APPLICATION NUMBER: US 07/971,160  
49 (B) FILING DATE: 05-NOV-1992  
50  
51 (vii) PRIOR APPLICATION DATA:  
52 (A) APPLICATION NUMBER: US 07/879,689  
53 (B) FILING DATE: 05-MAY-1992  
54  
55 (viii) ATTORNEY/AGENT INFORMATION:  
56 (A) NAME: Meyers, Kenneth J.  
57 (B) REGISTRATION NUMBER: 25,146  
58 (C) REFERENCE/DOCKET NUMBER: 3495-0111-11  
59  
60 (ix) TELECOMMUNICATION INFORMATION:  
61 (A) TELEPHONE: 202-408-4000  
62 (B) TELEFAX: 202-408-4400  
63  
64  
65 (2) INFORMATION FOR SEQ ID NO:1:  
66  
67 (i) SEQUENCE CHARACTERISTICS:  
68 (A) LENGTH: 714 base pairs  
69 (B) TYPE: nucleic acid  
70 (C) STRANDEDNESS: single  
71 (D) TOPOLOGY: linear  
72  
73 (ii) MOLECULE TYPE: DNA (genomic)  
74  
75  
76 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:  
77  
78 ATGCATATGA AAAACATCAA AAAAAACCAG GTAATGAACC TCGGTCCGAA CTCTAAACTG 60  
79 CTGAAAGAAT ACAAATCCCA GCTGATCGAA CTGAACATCG AACAGTTCGA AGCAGGTATC 120  
80 GGTCTGATCC TGGGTGATGC TTACATCCGT TCTCGTGATG AAGGTAAAAC CTACTGTATG 180  
81 CAGTTCGAGT GGAAAAACAA AGCATAATG GACCACGTAT GTCTGCTGTA CGATCAGTGG 240  
82 GTACTGTCCC CGCCGCACAA AAAAGAACGT GTTAACCACC TGGGTAACCT GGTAATCACC 300  
83 TGGGGCGCCC AGACTTTCAA ACACCAAGCT TTCAACAAAC TGGCTAACCT GTTCATCGTT 360  
84 AACAACAAAA AAACCATCCC GAACAACCTG GTTGAAAACT ACCTGACCCC GATGTCTCTG 420  
85 GCATACTGGT TCATGGATGA TGGTGGTAAA TGGGATTACA AAAAAAATC TACCAACAAA 480  
86 TCGATCGTAC TGAACACCCA GTCTTTCACT TTCGAAGAAG TAGAATACCT GGTTAAGGGT 540  
87 CTGCGTAACA AATTCCAACCT GAACTGTTAC GTAAAAATCA AAAAAACAA ACCGATCATC 600  
88 TACATCGATT CTATGTCTTA CCTGATCTTC TACAACCTGA TCAAACCGTA CCTGATCCCG 660  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99

RAW SEQUENCE LISTING  
PATENT APPLICATION US/09/492,697DATE: 02/16/2000  
TIME: 02:14:57

INPUT SET: S34756.raw

100 CAGATGATGT ACAAACTGCC GAACACTATC TCCTCCGAAA CTTTCCTGAA ATAA 714  
101  
102  
103 (2) INFORMATION FOR SEQ ID NO:2:  
104  
105 (i) SEQUENCE CHARACTERISTICS:  
106 (A) LENGTH: 237 amino acids  
107 (B) TYPE: amino acid  
108 (D) TOPOLOGY: linear  
109  
110 (ii) MOLECULE TYPE: peptide  
111  
112  
113 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:  
114  
115 Met His Met Lys Asn Ile Lys Lys Asn Gln Val Met Asn Leu Gly Pro 15  
116 1 5 10  
117  
118 Asn Ser Lys Leu Leu Lys Glu Tyr Lys Ser Gln Leu Ile Glu Leu Asn 30  
119 20 25  
120  
121 Ile Glu Gln Phe Glu Ala Gly Ile Gly Leu Ile Leu Gly Asp Ala Tyr 45  
122 35 40  
123  
124 Ile Arg Ser Arg Asp Glu Gly Lys Thr Tyr Cys Met Gln Phe Glu Trp 60  
125 50 55  
126  
127 Lys Asn Lys Ala Tyr Met Asp His Val Cys Leu Leu Tyr Asp Gln Trp 80  
128 65 70 75  
129  
130 Val Leu Ser Pro Pro His Lys Lys Glu Arg Val Asn His Leu Gly Asn 95  
131 85 90  
132  
133 Leu Val Ile Thr Trp Gly Ala Gln Thr Phe Lys His Gln Ala Phe Asn 110  
134 100 105  
135  
136 Lys Leu Ala Asn Leu Phe Ile Val Asn Asn Lys Lys Thr Ile Pro Asn 125  
137 115 120  
138  
139 Asn Leu Val Glu Asn Tyr Leu Thr Pro Met Ser Leu Ala Tyr Trp Phe 140  
140 130 135  
141  
142 Met Asp Asp Gly Gly Lys Trp Asp Tyr Asn Lys Asn Ser Thr Asn Lys 160  
143 145 150 155  
144  
145 Ser Ile Val Leu Asn Thr Gln Ser Phe Thr Phe Glu Glu Val Glu Tyr 175  
146 165 170  
147  
148 Leu Val Lys Gly Leu Arg Asn Lys Phe Gln Leu Asn Cys Tyr Val Lys 190  
149 180 185  
150  
151 Ile Asn Lys Asn Lys Pro Ile Ile Tyr Ile Asp Ser Met Ser Tyr Leu  
152

# RAW SEQUENCE LISTING PATENT APPLICATION US/09/492,697

DATE: 02/16/2000  
TIME: 02:14:58

INPUT SET: S34756.raw

```

153                                     195               200               205
154                                     Ile Phe Tyr Asn Leu Ile Lys Pro Tyr Leu Ile Pro Gln Met Met Tyr
155                                     210               215               220
156
157                                     Lys Leu Pro Asn Thr Ile Ser Ser Glu Thr Phe Leu Lys
158                                     225               230               235
159
160
161 (2) INFORMATION FOR SEQ ID NO:3:
162
163     (i) SEQUENCE CHARACTERISTICS:
164         (A) LENGTH: 722 base pairs
165         (B) TYPE: nucleic acid
166         (C) STRANDEDNESS: single
167         (D) TOPOLOGY: linear
168
169     (ii) MOLECULE TYPE: DNA (genomic)
170
171
172     (xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:
173
174     AAAAATAAAA TCATATGAAA AATATTAAAA AAAATCAAGT AATCAATCTC GGCCTATTT      60
175
176     CTAAATTATT AAAAGAATAT AAATCACAAT TAATTGAATT AAATATTGAA CAATTTGAAG      120
177
178     CAGGTATTGG TTTAATTTTA GGAGATGCTT ATATTCGTAG TCGTGATGAA GGTAAACTT      180
179
180     ATTGTATGCA ATTTGAGTGG AAAAATAAGG CATACATGGA TCATGTATGT TTATTATATG      240
181
182     ATCAATGGGT ATTATCACCT CCTCATAAAA AAGAAAGAGT TAATCATTTA GGTAATTTAG      300
183
184     TAATTACCTG GGGAGCTCAA ACTTTTAAAC ATCAAGCTTT TAATAAATTA GCTAACTTAT      360
185
186     TTATTGTAAA TAATAAAAAA CTTATTCCTA ATAATTTAGT TGAAAATTAT TTAACACCTA      420
187
188     TGAGTCTGGC ATATTGGTTT ATGGATGATG GAGGTAAATG GGATTATAAT AAAAATTCTC      480
189
190     TTAATAAAAAG TATTGTATTA AATACACAAA GTTTTACTTT TGAAGAAGTA GAATATTTAC      540
191
192     TTAAAGGTTT AAGAAATAAA TTTCAATTAA ATTGTTATGT TAAAATTAAT AAAAATAAAC      600
193
194     CAATTATTTA TATTGATTCT ATGAGTTATC TGATTTTTTA TAATTTAATT AAACCTTATT      660
195
196     TAATTCCTCA AATGATGTAT AAACGCTTA ATACTATTC ATCCGAAACT TTTTAAAT      720
197
198     AA
199
200
201 (2) INFORMATION FOR SEQ ID NO:4:
202
203     (i) SEQUENCE CHARACTERISTICS:
204         (A) LENGTH: 235 amino acids
205         (B) TYPE: amino acid

```

DATE: 02/16/2000  
TIME: 02:14:58

```

206          (D) TOPOLOGY: linear
207
208 (ii) MOLECULE TYPE: peptide
209
210
211 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:
212
213 Met Lys Asn Ile Lys Lys Asn Gln Val Met Asn Leu Gly Pro Asn Ser
214 1              5              10              15
215
216 Lys Leu Leu Lys Glu Tyr Lys Ser Gln Leu Ile Glu Leu Asn Ile Glu
217 20              25              30
218
219 Gln Phe Glu Ala Gly Ile Gly Leu Ile Leu Gly Asp Ala Tyr Ile Arg
220 35              40              45
221
222 Ser Arg Asp Glu Gly Lys Thr Tyr Cys Met Gln Phe Glu Trp Lys Asn
223 50              55              60
224
225 Lys Ala Tyr Met Asp His Val Cys Leu Leu Tyr Asp Gln Trp Val Leu
226 65              70              75              80
227
228 Ser Pro Pro His Lys Lys Glu Arg Val Asn His Leu Gly Asn Leu Val
229 85              90              95
230
231 Ile Thr Trp Gly Ala Gln Thr Phe Lys His Gln Ala Phe Asn Lys Leu
232 100             105             110
233
234 Ala Asn Leu Phe Ile Val Asn Asn Lys Lys Leu Ile Pro Asn Asn Leu
235 115             120             125
236
237 Val Glu Asn Tyr Leu Thr Pro Met Ser Leu Ala Tyr Trp Phe Met Asp
238 130             135             140
239
240 Asp Gly Gly Lys Trp Asp Tyr Asn Lys Asn Ser Leu Asn Lys Ser Ile
241 145             150             155             160
242
243 Val Leu Asn Thr Gln Ser Phe Thr Phe Glu Glu Val Cys Tyr Leu Val
244 165             170             175
245
246 Lys Gly Leu Arg Asn Lys Phe Gln Leu Asn Cys Tyr Val Lys Ile Asn
247 180             185             190
248
249 Lys Asn Lys Pro Ile Ile Tyr Ile Asp Ser Met Ser Tyr Leu Ile Phe
250

```

PAGE: 1

**SEQUENCE VERIFICATION REPORT**  
**PATENT APPLICATION US/09/492,697**

DATE: 02/16/2000  
TIME: 02:14:58

*INPUT SET: S34756.raw*

Line	Error	Original Text
31	Wrong application Serial Number	(A) APPLICATION NUMBER: Unassigned